云南植物研究 1990，12（2）：179—185
Acta Botanica Yunnanica

植物区系的数值分析

左家哺

（贵州省林业学校，修文县550201）

摘要 本文介绍了应用数值分析的某些方法和原理进行植物区系基本成分统计、地理成分划分及其分区。植物区系数值分析不仅具有分析严密、计算简单及结果直观等优点，而且它比经典方法更能反映其本质特征。

关键词 植物区系；数值分类学

A NUMERICAL ANALYSIS OF FLORA

Zuo Jiafu

(*Guizhou Forestry School, Xiuwen county*550201)

Abstract This paper introduces some methods and principles of the numerical analysis that has been used for the statistics of basic elements, division of geographical elements and regionalization of flora. The methods of the numerical analysis of flora have the merits of strictness of analysis, simplicity of calculation and intuition of result, and that can reflect more the intrinsic character of flora than the classical methods.

Key words Flora; Numerical taxonomy

以往，大多数学者研究植物区系都是沿用经典分析方法[1]，难免或多或少带有人为的主观意识。因此，导致研究同一个问题得出的结论不一样，甚至分歧较大[2－6]。其主要原因除了各个学者所掌握的材料、研究环境及探究问题的角度等不同外，还与经典方法本身存在某些缺陷不无关系。这样一来，势必要寻求研究植物区系的新方法。

数值分类学（Numerical Taxonomy）自本世纪50年代诞生以来，在生物学各个领域如分类、生态、生物地理、生物化学、古生物、遗传、进化及医学等学科中广泛应用。在植物区系方面，王荷生教授研究了我国种子植物特有属[7]、徐炳声等人研究了我国荚蒾属（*Viburnum*）的分布式样[8]，笔者在研究贵州润楠属（*Machilus*）1)、含笑

1989-04-12收稿

属（*Michelia*）[9]2）、猕猴桃属（*Actinidia*）[10]、槭树科（Aceraceae）3）等类群植物区系工作中运用数值分类学的某些原理和方法[11]取得了比较理想的结果，比经典方法更能反映其本质特征。

植物区系的数值分析（Floristic Numerical Analysis）是将植物区系诸定性内容转化为定量，运用计算机助于概率论、统计学、信息论、模糊数学、图论和拓扑等数学方法进行研究。本文就植物区系成分统计、区系成分划分及植物区系分区的数值分析进行扼要的阐述，以期抛砖引玉，使之日臻完善。

基本成分的数值统计分析

1.植物各类单位的统计

一个地区植物各级分类单位的统计是植物区系分析的最基本问题，它能表明该地区植物区系的丰富程度。如中国种子植物有301科2980属24550种[3]，云南有240科1984属13000种[12]；樟科（Lauraceae）全球有45属2000—2500种，中国有20属471种（含种下等级）[13]等等。

2.多个地区植物区系成分的综合比较分析

单独统计一个地区植物区系成分的意义不是很大，只有通过多个地区的综合比较才更说明问题。对此，以往都是通过列表（表1）来描述。如果被比较的地区非常之多，而且这些地区植物科、属、种的数量不相上下时，这种列表法就难于奏效。

表1 中国若干省区种子植物区系成分的比较[12，14—20]4）

Table 1 Comparison in the floristic elements of seed plants among some provinces or regions in China

分类单位	贵州	云南	四川	广西	湖南	湖北	广东	安徽	海南
科	198	240	119	235	203	172	232	155	203
属	4114	1984	4101	1545	1113	1098	1575	935	1116
种	4951	13000	8546	5443	3681	3876	6063	3020	2860

因此，本文给出一种综合比较的简捷方法：

$$S_j=\sum_{j=1}^{n}(X_{ij}-\overline{X}_{ij}/\overline{X}_{ij})$$

式中：1）、X_{ij}表示K个地区中第i个地区n个分类单位中第j个分类单位的数

1）左家哺，贵州润楠属植物区系与分布（手稿）。

2）左家哺，应用模糊聚类分析划分贵州含笑属植物区系地理成分，云南植物研究（待刊）。

3）左家哺，贵州槭树科植物区系与分布的初步研究，武汉植物学研究（待刊）。

值；

2）$\overline{X}_{ij}=\frac{1}{K}\sum_{j=1}^{K}X_{ij}$表示K个地区中n个分类单位中第j个分类单位的平均值；

3）S_i表示K个地区中第i个地区植物区系成分的综合系数（integrative coefficient）。S_i越大，第i个地区植物区系越丰富；相反，则越贫乏。

以S_1公式综合比较分析多个地区植物区系成分丰富程度的方法称为综合系数法（Method of the integrative coefficients）。以表1为例，这9个省区种子植物区系成分的S_i值是：云南为1.921、四川为0.470、广东为0.367、广西为0.250、贵州为－0.115、湖南为－0.544、湖北为－0.665、海南为－0.685、安徽为－1.018。由此表明，这9个省区种子植物区系丰富程度依次递减，其中云南最丰富，安徽最贫乏。不难看出，本方法具有计算简单、结果直观、一目了然等优点。

5.植物区系成分相似性分析

在多个地区植物区系成分相似性分析中，D. Szymkiewicz 1934年首先将属相似性指标引入地中海周围地区植物区系研究中，其方法是：某个地区与另一个地区共有属数与该地区总属数（两者都不含世界分布属）的比值再乘100，通过研究认为两地属相似性指标为0.5(或50%)以上时，为近亲植物区系；相反，近亲性不大或完全不存在[1]。接着，S. Kulczunski 1944年又确定了科相似性系数，其方法是：某个地区同印度支那共有科数除以一切科的数目再乘100。由此，他绘制了地球上植物区系科相似性系数等值线图[1]。

作者在研究贵州润楠属1）、含笑属[9]、猕猴桃属[10]和槭树科3）时，采用关联系数的几个公式[11]：

（1）$a/a+b+c$

（2）$2a/2a+b+c$

（3）$a/\sqrt{(a+b)(b+c)}$

式中：a表示两个分布地区共有的种数；b表示甲分布地区存在而乙分布地区不分布的种数；c表示乙分布地区存在而甲分布地区不分布的种数

分析了贵州与邻近地区植物区系种的相似性系数(表2)。结果表明，贵州与广西、湖南这几个类群的种相似性系数比云南、四川等地都要大些，也就是说，贵州与广西、湖南的植物区系比其它地区更亲近、差异更小。这一问题似乎不好理解，实质上，它反映了贵州这些类群在早期演化、发展方面与广西、湖南有着密切的关系。从古地理背景亦可得到佐证，早奥陶纪到晚三叠纪，贵州东北部、东南部和南部（稍偏东）与湖南、广西

4）贵州资料取自《贵州种子植物分类、分布与用途》（黄威廉等，1983）；广西资料取自《广西植物名录》（1～3册）（广西植物所，1970—1973）。

（部分地区）为江南古陆之一部分，其余地区为海水所侵，晚侏罗纪该古陆西南边缘才扩展到贵阳至昆明一线，直至早白垩纪[21]。众所周知，目前许多学者都认为有花植物在白垩纪或晚侏罗纪才产生（E. B. 吴鲁夫，1932；J. Hutchinson，1926；A. L. Takhtajian，1969；田村道夫，1974；C. R. Soo，1775；K.R.Sporne，1976；H. Gaussen，1961；A. Cronquist；1978；路安民等，1978；R. Dahlgren，1983；徐仁，1987；等等）。然而，根据大陆漂移理论，联合古陆在三叠纪至侏罗纪开始自东部出现分裂，产生古地中海，晚侏罗纪或早白垩纪已完全分开，形成劳亚、冈瓦纳两大古陆[6]。倘若按上述提到的有花植物起源时间，那么现代有花植物全球性遍布，并且各个大陆之间植物相互联系及共有性就难于解释。因此，有花植物肯定在联合古陆解体之前就已产生，并广布全球，至少各主要科、属是如此。综合看来，贵州这些类群与广西、湖南的关系较云南、四川亲近是在地史时期形成的，且它们的早期演化、发生发展密切相关。

表2 贵州与邻近地区几个分类群的种相似性系数

Table 2 The similarity coefficients of the species of the several taxa of Guizhou with adjacent regions

类群	广西	湖南	广东	云南	四川	江西	湖北
润楠属 *Machilus*	0.51	0.35	0.35	0.33	0.18	0.16	0.11
含笑属 *Michelia*	0.71	0.59	0.53	0.38	0.25	0.27	0.25
猕猴桃属 *Actinidia*	0.36	0.42	0.24	0.33	0.33	0.29	0.31
槭树科 *Aceraceae*	0.52	0.46	0.37	0.23	0.34	0.29	0.44

植物区系成分相似性分析不但可以得到多个地区之间植物区系亲近程度情况以及帮助推断它们的起源，而且对于划分植物区系成分及其植物区系分区也起着十分重要的作用。

目前，相似性系数计算方法有十余种，如P. Sneath和R.Sokal 1973年合著的《数值分类学——数值分类学的原理和应用》[22]和阳含熙、卢泽愚1981年出版的《植物生态学的数量分类方法》[11]两书中详细介绍了距离系数、结合系数（或称关联系数）、相关系数（或称内积系数）和概率相似系数（或称信息系数）等四大类相似性系数的公式及其计算。

此外，笔者在分析了贵州几个类群植物区系成分的种相似性系数之后，认为王荷生教授所拟定的等级[7]亦适合种间相似性分析，现摘录如表3。

植物区系成分的划分

一个植物区系按其起源、发生发展或植物地理等规律可以划分为若干群，每一群称为一类区系成分，其类型有地理成分、发生成分、历史成分、迁移成分和生态成分等5种[23]。其中，地理成分被广泛应用而且比较容易掌握，它是以现代植物分布区为基础

而划分的一类区系成分。

表 3 种相似性系数的等级[7]

Table 3 The grades of similarity coefficient of the species

相似性系数（Sc）	亲近程度
Sc⩾0.50	很亲近
0.3⩽Sc<0.5	亲　近
0.2⩽Sc<0.3	比较亲近
0.1⩽Sc<0.2	比较疏远
0.05⩽Sc<0.1	疏　远
0.01⩽Sc<0.05	很疏远
Sc＝0	无关系

划分地理成分可以通过聚类分析来实现。目前，聚类分析一般有系统聚类法、图论聚类法、主分量聚类法、信息聚类法和模糊聚类法等 5 种[11, 12]。笔者曾应用系统、信息和模糊等 3 种方法研究过植物区系地理成分的划分，都取得了比较满意的结果。

本文以模糊聚类分析划分植物区系地理成分为例，简单介绍其主要步骤和方法。

1.收集植物地理分布资料建立其模糊矩阵

对研究范围内所有植物的地理分布资料，以某一种类在某分布区有分布记为 1，相反，则记为 0，列出植物地理分布数据表即构成了一个模糊矩阵。

2.建立模糊相似关系及其矩阵

对于建立两个待划分的成分之间模糊相似关系的方法常有13种[24]。笔者以为对于植物地理分布二元状态数据而言，以笔者所推导的信息系数减数法公式[2)]：

$$r_{ij}=\begin{cases}1; & \text{当 } i=j \text{ 时}\\ 1-Mt_{ij}; & \text{当 } i\neq j \text{ 时}\end{cases}$$

式中：1）$M=c\ 2\ln 2$；c 取适当的值时使 $0\leqslant r_{ij}\leqslant 1$。

2）t_{ij}表示两类群仅分布其中一个类群的分布地区数；如果在研究植物区系分区时，则表示两个分布地区仅出现在其中一分布地区的类群数。

来建立其模糊相似关系及其矩阵比其它公式更符合实际，也更简单。

3.建立模糊等价矩阵

一般而言，模糊相似关系难于满足传递性，因此采用求包含模糊相似关系矩阵的最小传递矩阵的方法来构造一个模糊等价矩阵，直接可以用该矩阵来对全体成分进行划分。

4.归类处理

虽然能以模糊等价矩阵对待划分的全体成分实现划分植物区系地理成分，但是比较麻烦。若以

$$r_{ij}^{(\lambda)}=\begin{cases}1; & 当 r_{ij}\geqslant\lambda 时\\ 0; & 当 r_{ij}<\lambda 时\end{cases}$$

对模糊等价矩阵中的元素进行变换。那么，由$r_{ij}^{(\lambda)}$为元素构成的矩阵称为λ-截矩阵（λ为阈值或置信水平），它表示一种普通等价关系；加之，$0\leqslant\lambda<r_{ij}\leqslant 1$，故按模糊等价矩阵对全体成分分成的每一类必定量按$\lambda$-截矩阵将全体成分划分的类的子类。

λ的每一次取值，全体成分中某些成分被划分为一类或几类，直至最小λ值。同时，由λ越来越小时，划分的区系成分越来越粗放，逐步形成一类。这个过程中形成了一个动态树状图（或聚类图）。通过贵州含笑属的研究表明，贵州所产17种含笑可以划分为7种类型的地理成分2），与常规的经典分析的结果基本相符[9]。不仅如此，以模糊等价关系划分植物区系地理成分类型的指标比经典方法更能反映植物区系的本质特征，也能准确地反映各类地理成分之间的关系。这些是经典方法做不到的。

以上介绍的是求模糊相似矩阵的传递闭包法，如果待划分的成分较多时，这种方法十分麻烦，那么采用求最大树法比较简便。其方法是：先画出被分配成分的集合，从模糊相似矩阵中按r_{ij}的大小顺序用直线连接起来，并标明权重。若在某一步出现回路，便不画这一步，直到所有元素连通为止。这样就得到1棵所谓的最大树（虽然最大树不止1棵，但可证明并不影响聚类的结果）。取定λ，的去掉权重低于λ的连线，即可将元素（或成分）划分，互相连通的元素（或成分）被归为同一类[24]。笔者采用最大树法在研究中国油杉属（*Keteleeria*）分布式样时9），同样取得了令人满意的效果。

植物区系的分区

目前，大部分学者主要根据研究范围内各地的植物区系起源、发生发展、特有成分（科、属、种）数量以及各地植被的区系组成等情况来进行植物区系的分区。植物区系的分区亦可通过聚类分析来实现，即选取在研究范围内广泛分布的几个代表类群，对研究范围划分为若干个小区进行聚类，并辅以各小区的植被情况，种类分布特征与生态地理等综合分析，将会得到既符合实际又能指导生产工作的植物分区结论。

笔者运用信息聚类方法分析贵州含笑属植物的分布式样时[9]，将江南划分为华南、华东、华中、西南、海南等5个地区，这与吴征镒教授的研究结论基本一致[2]。徐炳声教授等人通过中国荚蒾属植物分布式样研究[8]，将我国划分为南方、北方两个植物区，以下归为8个亚区即东北、西北、中南、台湾、华南、西南、横断山脉和藏西等，这也与吴征镒教授的中国植物区系分区中的区、亚区基本相符[2]。

9）左家哺，中国油杉属分布式样的模糊聚类分析，中南林学院学报（待刊）。

结　束　语

植物区系的数值分析与经典方法相比，不仅具有计算一次完成、结果直观简单，很多方法具有严格的逻辑性和严密的数学证明，而且还能更准确地反映植物区系的本质特征及其各种植物区系之间的亲密或疏远关系。这样一来，植物区系的数值分析只要得到重视亦会象数值分类学一样迅猛发展、完善。

致谢　本文承中国科学院昆明植物所吴征镒教授修改、审阅。

参　考　文　献

1　Szafer W．（傅子祯译）．普通植物地理学原理．北京：高等教育出版社，1958:351—359
2　吴征镒．云南植物研究1979　1（1）：1—22
3　吴征镒，王荷生．中国自然地理，植物地理（上）．北京：科学出版社，1983:102—103
4　张宏达．中山大学学报（自然科学版）　1980（1）:89—98
5　张宏达．中山大学学报（自然科学版）　1984（4）:93—101
6　张宏达．中山大学学报（自然科学版）　1986（3）：1—11
7　王荷生．植物分类学报　1985；23（4）:241—258
8　徐炳声，缪柏茂．植物分类学报　1988；26（5）:329—342
9　左家哺．黔南林业科技　1988（1）：1—10
10　左家哺．经济林研究　1988；6（2）:20—23
11　阳含熙，卢泽愚．植物生态学的数量分类方法．北京：科学出版社，1981：33—64、90—120
12　李锡文．云南植物研究　1985；7（4）:361—382
13　李锡文等．中国植物志（31卷）　北京：科学出版社，1984：2
14　《四川植被》协作组．四川植被．成都：四川人民出版社，1980:50
15　祁承经．植物研究　1984；4（1）:130—145
16　郑重．武汉植物学研究　1983；1（2）:165—188
17　广东植物所．广东植被．北京：科学出版社　1976:18—21
18　《安徽植被》协作组．安徽植被．合肥：安徽科学技术出版社　1983:26
19　张超常等．中山大学学报（自然科学版）　1983（3）:53—67
20　缪汝槐．中山大学学报（自然科学版）　1986（4）:57—63
21　宋春青，张振春．地质学基础．北京：人民教育出版社，1978:508—555
22　Sneath P.和Sokal R．（赵铁桥译，汪振儒校）．数值分类学——数值分类学的原理和应用．北京：科学出版社．1984:76—206
23　吴鲁夫 E．B．（仲崇信、张梦庄译）．历史植物地理学引论．北京：科学出版社，1960：226—227
24　葛苏林．模糊子集·模糊关系·模糊映射．北京：北京师范大学出版社，1985：72—75、85—86